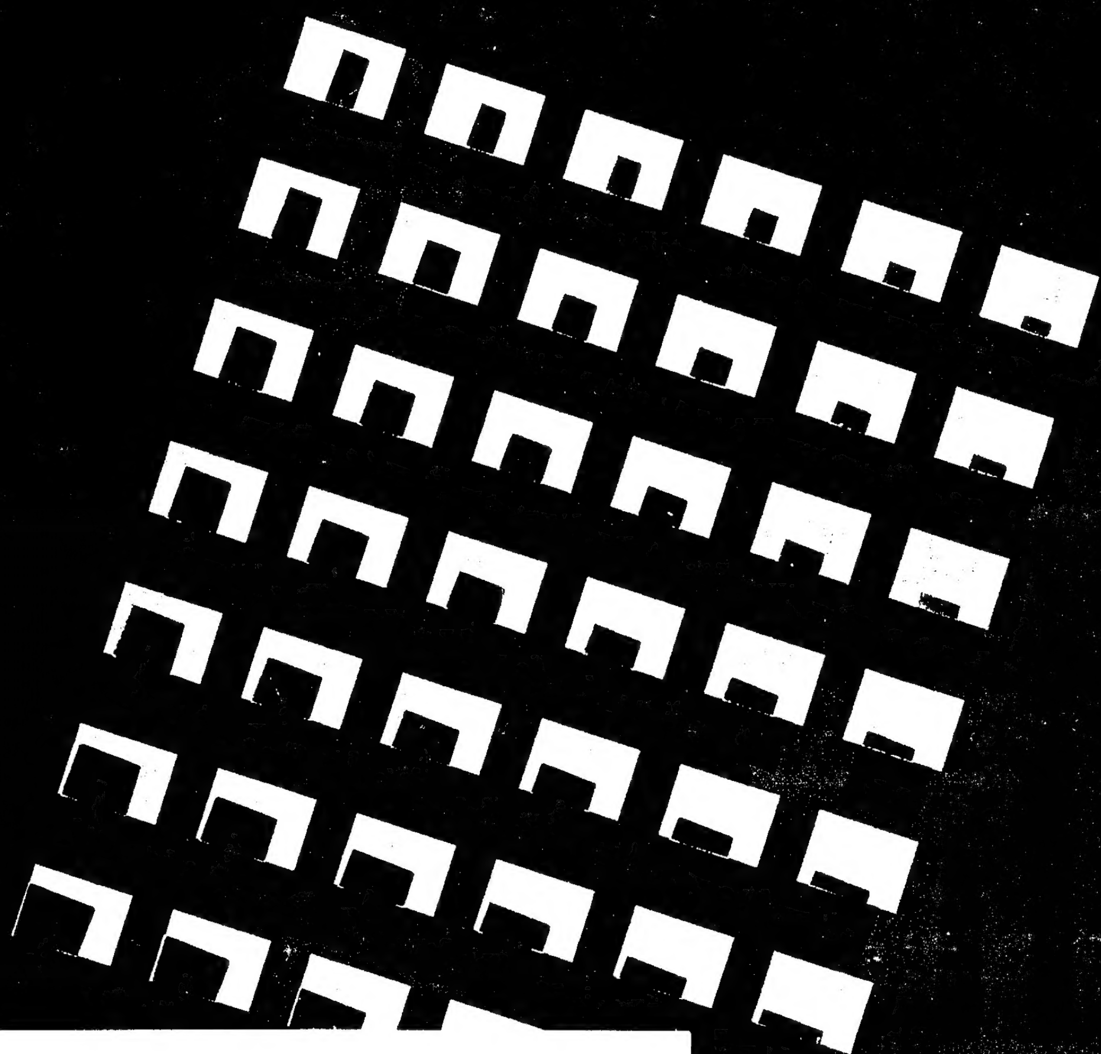
TNO-report
TM-96-A026

title
Physiological workload reactions to increasing levels of task difficulty

TNO Human Factors
Research Institute



TNO-report
TM-96-A026

TNO Human Factors
Research Institute

Kampweg 5
P.O. Box 23
3769 ZG Soesterberg
The Netherlands

Phone +31 346 35 62 11
Fax +31 346 35 39 77

title
Physiological workload reactions to increasing levels of task difficulty

authors
J.A. Veltman
A.W.K. Gaillard

date
21 June 1996

number of pages : 28 (incl. appendices, excl. distribution list)

TNO Human Factors Research Institute
is part of TNO Defence research
to which also belong:
TNO Physics and Electronics Laboratory
TNO Prins Maurits Laboratory

Netherlands Organization for
Applied Scientific Research

REPORT DOCUMENTATION PAGE

1. DEFENCE REPORT NUMBER (MOD-NL) RP 96-0162	2. RECIPIENT'S ACCESSION NUMBER	3. PERFORMING ORGANIZATION REPORT NUMBER TM-96-A026
4. PROJECT/TASK/WORK UNIT NO. 789.4	5. CONTRACT NUMBER A91/KLu/314	6. REPORT DATE 21 June 1996
7. NUMBER OF PAGES 28	8. NUMBER OF REFERENCES 28	9. TYPE OF REPORT AND DATES COVERED Final

10. TITLE AND SUBTITLE

Physiological workload reactions to increasing levels of task difficulty

11. AUTHOR(S)

J.A. Veltman and A.W.K. Gaillard

12. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

TNO Human Factors Research Institute
Kampweg 5
3769 DE SOESTERBERG

13. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Director of Airforce Research and Development
Binckhorstlaan 135
2516 BA DEN HAAG

14. SUPPLEMENTARY NOTES

15. ABSTRACT (MAXIMUM 200 WORDS, 1044 BYTE)

The sensitivity of physiological measures to mental workload has been investigated in a flight simulator. Twelve pilots had to fly through a tunnel with different levels of difficulty. Additionally, they had to perform a memory task with four levels of difficulty. The easiest memory task was combined with the easiest tunnel task and the most difficult memory task with the most difficult tunnel task. Between each tunnel session, subjects had to fly a pursuit task in which a target jet had to be followed at a large distance. Rest periods before and after the experiment were used as a baseline for the physiological measures. Mental workload was measured with heart period, continuous blood pressure, respiration and eye blinks. Heart rate variability, blood pressure variability and the gain between systolic blood pressure and heart period (modulus) was derived from the measures. All measures showed differences between rest and flight, and between the pursuit task and the tunnel task. Only heart period was sensitive to the different levels of the tunnel. Heart rate variability is affected by respiration, and therefore, can only be interpreted together with the respiratory data. The modulus was hardly influenced by respiration and therefore, appears to be a better measure than heart rate variability. From the respiratory parameters, the duration of a respiratory cycle was the most sensitive to changes in workload. The time in between two successive eye blinks (blink interval) increased and the blink duration decreased as more visual information had to be processed. Increasing the difficulty of the memory task leaded to a decrement in blink interval, probably caused by subvocal activity during the rehearsal of the letters. Blink duration was not influenced by the memory load.

16. DESCRIPTORS	IDENTIFIERS
Blood Pressure Eye Blinks Heart Rate Pilot Psychophysiology Respiration Workload	

17a. SECURITY CLASSIFICATION (OF REPORT)	17b. SECURITY CLASSIFICATION (OF PAGE)	17c. SECURITY CLASSIFICATION (OF ABSTRACT)
18. DISTRIBUTION/AVAILABILITY STATEMENT Unlimited availability		17d. SECURITY CLASSIFICATION (OF TITLES)

TNO Technische Menskunde

Kampweg 5
Postbus 23
3769 ZG Soesterberg

Telefoon 0346 356 211
Fax 0346 353 977

Drs. E.A. van Hoek
Directeur Wetenschappelijk
Onderzoek en Ontwikkeling
Ministerie van Defensie
Directoraat-Generaal Materieel
Postbus 20701
2500 ES DEN HAAG

Doorkiesnummer
0346 356 245

Datum
27 juni 1996

Ons nummer
96 rap 072

Uw brief

Onderwerp
Rapport TM-96-A026

Hierbij hebben wij het genoegen u het volgende rapport toe te zenden:

Rapport TM-96-A026: "Physiological workload reactions to increasing levels of task difficulty", door Drs. J.A. Veltman en Prof.dr. A.W.K. Gaillard.

Het onderzoek werd uitgevoerd in het kader van opdrachtnummer A91/KLu/314.

Dr.ir. A. van Meeteren,
directeur

Bijlage(n)

mt

Managementuittreksel TNO Technische Menskunde, Soesterberg

titel	:	Physiological reactions to increasing levels of task difficulty (Fysiologische reacties bij toenemende niveaus van taakmoeilijkheid)
auteurs	:	Drs. J.A. Veltman en Prof. dr. A.W.K. Gaillard
datum	:	21 juni 1996
opdrachtnr.	:	A91/KLu/314
IWP-nr.	:	789.4
rapportnr.	:	TM-96-A026

Dit onderzoek is uitgevoerd binnen de raamopdracht "werklast vliegers" (A89/KLu/322) dat moet resulteren in een bruikbaar concept voor mentale werkbelasting en een set van maten waarmee in operationele situaties de werkbelasting kan worden bepaald. In voorgaande experimenten (Veltman & Gaillard, 1993, 1994, 1995) is de gevoeligheid van fysiologische werkbelastingsmaten bekeken bij grote verschillen in taakmoeilijkheid. In het huidige experiment is de gevoeligheid voor kleine verschillen in taakmoeilijkheid nagegaan.
Twaalf proefpersonen (vliegers) moesten, in de vliegsimulator van TNO-TM, door een tunnel vliegen en gelijktijdig een geheugentaak uitvoeren. De moeilijkheid van de tunneltaak en de geheugentaak werden gelijktijdig gevarieerd in vier niveaus. Voor en na elke tunneltaak voerden de proefpersonen een makkelijke volgtaak uit, waarbij een doelvliegtuig op grote afstand moest worden gevolgd. Een belangrijk aspect van de mentale werkbelasting is de inspanning die moet worden geleverd om de taak uit te kunnen voeren. Deze inspanning werd gemeten met behulp van fysiologische maten. Fysiologische metingen tijdens een rustperiode voor en na het experiment werden gebruikt als referentie. Mentale inspanning werd gemeten met behulp van hartslaginterval, continue bloeddruk, ademhaling en oogknipperingen. Daarnaast werden de hartslagvariabiliteit, bloeddrukvariabiliteit en de overdrachtsfunctie tussen systolische bloeddruk en hartslag (modulus) afgeleid. Alle maten waren gevoelig voor de verschillen tussen rust en vliegen en tussen de volg- en tunneltaak. Alleen hartslaginterval was gevoelig voor de verschillen in moeilijkheidsniveaus in de tunnels. Hartslagvariabiliteit werd beïnvloed door de ademhaling waardoor deze maat alleen goed geïnterpreteerd kan worden als ook de ademhalingsgegevens bekend zijn. De modulus werd veel minder beïnvloed door de ademhaling en kan daardoor zonder gegevens over de ademhaling worden geïnterpreteerd. Ten gevolge van de visuele belasting werd er minder en sneller met de ogen geknipperd. Bij toenemende geheugenbelasting werd er meer geknipperd, terwijl de duur van de knipper gelijk bleef.

CONTENTS Page

Report No.:	TM-96-A026
Title:	Physiological workload reactions to increasing levels of task difficulty
Authors:	Drs. J.A. Veltman and Prof. Dr. A.W.K. Gaillard
Institute:	TNO Human Factors Research Institute Group: Work Environment
Date:	June 1996
DO Assignment No.:	A91/KLu/314
No. in Program of Work:	789.4

SUMMARY

The sensitivity of physiological measures to mental workload has been investigated in a flight simulator. Twelve pilots had to fly through a tunnel with different levels of difficulty. Additionally, they had to perform a memory task with four levels of difficulty. The easiest memory task was combined with the easiest tunnel task and the most difficult memory task with the most difficult tunnel task. Between each tunnel session, subjects had to fly a pursuit task in which a target jet had to be followed at a large distance. Rest periods before and after the experiment were used as a baseline for the physiological measures. Mental workload was measured with heart period, continuous blood pressure, respiration and eye blinks. Heart rate variability, blood pressure variability and the gain between systolic blood pressure and heart period (modulus) was derived from the measures. All measures showed differences between rest and flight, and between the pursuit task and the tunnel task. Only heart period was sensitive to the different levels of the tunnel. Heart rate variability is affected by respiration, and therefore, can only be interpreted together with the respiratory data. The modulus was hardly influenced by respiration and therefore, appears to be a better measure than heart rate variability. From the respiratory parameters, the duration of a respiratory cycle was the most sensitive to changes in workload. The time in between two successive eye blinks (blink interval) increased and the blink duration decreased as more visual information had to be processed. Increasing the difficulty of the memory task leaded to a decrement in blink interval, probably caused by subvocal activity during the rehearsal of the letters. Blink duration was not influenced by the memory load.

Rap.nr. TM-96-A026

TNO Technische Menskunde
Soesterberg

Fysiologische reacties bij toenemende niveaus van taakmoeilijkheid

J.A. Veltman en A.W.K. Gaillard

SAMENVATTING

De gevoeligheid van fysiologische maten voor mentale werkbelasting is onderzocht. Twaalf vliegers moesten, in de vliegsimulator van TNO-TM, door een tunnel vliegen en gelijktijdig een geheugentaak uitvoeren. Zowel de moeilijkheid van de tunneltaak als de geheugentaak werden gelijktijdig gevarieerd in vier niveaus. Voor en na elke tunneltaak voerden de proefpersonen een makkelijke volgtaak uit, waarbij een doelvliegtuig op grote afstand moest worden gevolgd. Een belangrijk aspect van de mentale werkbelasting is de inspanning die moet worden geleverd om de taak uit te kunnen voeren. Deze inspanning werd gemeten met behulp van fysiologische maten. Fysiologische metingen tijdens een rust periode voor en na het experiment werden gebruikt als referentie. Mentale inspanning werd gemeten met behulp van hartslaginterval, continue bloeddruk, ademhaling en oogknipperingen. Daarnaast werden de hartslagvariabiliteit, bloeddrukvariabiliteit en de overdrachtsfunctie tussen systolische bloeddruk en hartslag (modulus) afgeleid. Alle maten waren gevoelig voor de verschillen tussen rust en vliegen en tussen de volg- en tunneltaak. Alleen hartslaginterval was gevoelig voor de verschillen tussen de tunnels. Hartslagvariabiliteit werd beïnvloed door de ademhaling waardoor deze maat alleen goed geïnterpreteerd kan worden als ook de ademhalingsgegevens bekend zijn. De modulus werd veel minder beïnvloed door de ademhaling en kan daardoor zonder gegevens over de ademhaling worden geïnterpreteerd. Van de ademhalingsparameters was de ademhalingsduur het meest gevoelig voor veranderingen in de inspanning. Ten gevolge van de visuele belasting werd er minder en sneller met de ogen geknipperd. Bij toenemende geheugenbelasting werd er meer geknipperd, mogelijk ten gevolge van het subvocaal herhalen van de letters. De duur van de knipper werd niet beïnvloed door de geheugenbelasting.

1 INTRODUCTION

In complex systems, such as modern fighter jets and helicopters, operators have to manage several tasks at the same time; this has consequences for their mental workload. When new systems are introduced, it is important to evaluate the effects on mental workload. Before the workload can be evaluated it is important to know what factors can influence the workload and how workload can be measured. The present study is part of a project (under contract of the Netherlands Airforce) of which the aim is to gain insight into the concept of mental workload and to develop measurement techniques for mental workload. Before the present study is outlined, a short overview of mental workload will be presented.

1.1 Concept of mental workload

A generic definition of mental workload that is often used is that of Kantowitz (1988): *Workload is an intervening variable, similar to attention, that modulates or indexes the tuning between the demands of the environment and the capacity of the operator.* The workload is high when the difference between the demands and the available capacity is small. To measure mental workload it is necessary to know something about the task demands that influence workload and about the capacity of the operator. Because both operator capacity and task demand are difficult to measure, there are many operational definitions of mental workload. For example, Stone, Gulick and Gabriel (1987) define workload as the ratio of the available and required *time*. Workload is high when there is more time needed than available. Although this definition is only about one aspect of workload (time pressure), it can be used directly. Another operational definition is that of Hart and Staveland (1988) who define workload in terms of *costs* that are incurred by operators to achieve a particular level of performance. When it is important to have information about the causes of mental workload, information is needed about the capacity of the operator and about the task. The capacity is determined by stable factors like level of training, skills etc. and variable factors like noise, vibrations, fatigue, motivation, spatial awareness etc. Task demands are for a great part determined by the complexity of the task and the amount of attention required to perform the task (see Veltman, 1991). According to Hart and Staveland (1988), mental workload is determined by many aspects. They originally distracted 19 factors of which 6 were found to have a unique contribution to the overall workload: 1) mental demand—level of mental and perceptual activity required, 2) physical demand—level of physical activity required, 3) temporal demand—level of time pressure, 4) performance—satisfaction about own performance 5) effort—how hard someone has to work, mentally and physically, 6) frustration level—discouraged, irritated annoyed, etc.).

Operators can use different strategies to cope with workload. Hart (1989) showed that experienced operators work in advance during periods of low workload in order to eliminate workload peaks in the future. Hockey (1993) describes different strategies in a regulation model. Operators constantly compare their performance with the goal state. If the quality of the performance is not good enough according to the goal state, more effort will be invested. To a certain level this is an automatic process. An effort monitor evaluates the amount of effort that is required and when the effort increases too much, the performance evaluation

process is controlled at a higher cognitive level. Operators can apply different strategies to situations in which the performance level does not fit the goal state. They can decide to invest more effort or to decrease the goal and accept a lower level of performance. When these strategies are not possible because the performance level is already low or the operator has already invested a maximum amount of effort, the situation leads to stress. Thus, high levels of workload can lead to stress, but stress is not the same as high workload. As long as the operator has the capacity to invest the effort, performing the task can even give satisfaction. Gaillard and Wientjes (1994) show that there are substantial costs involved when one has to invest a lot of mental effort to perform a task that is mentally highly demanding. These costs are reflected in several physiological reactions.

The methods to measure workload are divided mostly into four groups (O'Donnell & Eggemeier, 1986):
Performance on the main task. In real life tasks, performance measures are often difficult to obtain due to technical reasons and can therefore not be used as a workload measure. Furthermore, performance should always be used in combination with other measures. For example, the performance can be low in an easy task because subjects do not invest effort in the task and it can be high in a difficult one because subjects invest a lot of effort. Thus, the level of performance without other information does not tell us anything about the workload.
Secondary task. The idea behind this measure is that when subjects have plenty of capacity left, the performance of a secondary task (which mostly has nothing to do with the main task) will be high and *vice versa.* A serious problem with these measures is that they may interfere with the main task. Subjects may lower the performance on the main task, in favour of the secondary task, which can lead to errors. In many situations this is not acceptable.
Subjective measures. Using these measures, subjects have to give ratings on one or several aspects of workload. They can be applied easily. However, they can be applied mostly only after the completion of a task, they do not discriminate between the effort invested by the subject and the difficulty of the task: small peaks in the workload will be overestimated, high workload at the beginning of a task will be underestimated and high workload at the end of a task will be overestimated.
Physiological reactions. Mental effort is reflected by several physiological changes. These measures have the advantage that they are objective and provide continuous information. Some disadvantages of physiological measures are: 1) they can be influenced by other factors like physical activity, 2) it is not easy to obtain data free of artifacts, 3) data analysis is time consuming. However, with the techniques that are available nowadays, the latter two points are becoming less important.

As described above, workload is influenced by many different factors and each measure reflects different aspects of workload. Therefore, workload can best be measured with a combination of measures. These measures can provide supplementary information about the level of workload in a particular task situation. For a complex task in which performance is difficult to measure, a combination of subjective and physiological techniques will give sufficient information about the workload. There are multidimensional subjective measures that require several aspects of workload to be rated and unidimensional measures that require a rating of the total workload or the effort invested in the task. It has been shown

that a unidimensional rating scale provides the same information in most cases as a multidimensional scale (Veltman & Gaillard, 1994; Hendy, Hamilton & Landry, 1993). Physiological measures reflect the costs that are involved when operators have to invest much effort. Because effort is only one aspect of workload, it is better to speak of physiological measures as a measure of mental effort than as a workload measure.

To evaluate physiological measures, two points are important: *motivation* and *task characteristics*. Subjects should be motivated to increase their effort whenever the task becomes more difficult, otherwise no changes in physiological measures will be found. Even when subjects are maximally motivated they will not invest all their effort in an easy task. Operators strive for an optimal level of performance instead of a maximal level. When the gain in performance is small related to the invested effort, operators will not invest more effort (Adams, Tenney & Pew, 1991). To vary the effort, the task should force the subjects to increase their effort. Not all tasks have these characteristics. Norman and Bobrov (1975) distinguished between two types of tasks: data- and resource-limited. In data-limited tasks, performance is determined by the quality of the incoming information. If the quality of this information is poor, investing more effort will have a negligible improvement in performance and, hence, subjects will not increase their effort. In resource-limited tasks, performance is dependent on the mental resources invested in the task. When such tasks become more difficult, increasing the effort will result in an increase in performance. In the previous study in which three different radars were compared, large differences in performance and subjective ratings were found, although the physiological reactions did not differ. The differences between these tasks were data-limited (the performance differences were caused by the qualities of the radars) and therefore no differences in physiological reactions were found.

1.2 Present study

The aim of the present experiment is to evaluate the sensitivity of physiological measures and their mutual relations. This is the fourth experiment in the workload project (A91/KLu/314). In the first two experiments, the task demands during flight were varied in two levels in separate blocks lasting about four minutes (Veltman & Gaillard, 1993, 1994). The task demands in the third experiment (Veltman, Gaillard & Van Breda, 1995) also had two levels, but the duration of the tasks was rather short (15 to 60 s). In each of these experiments, physiological measures discriminated between the two levels of difficulty. The parameters that were most sensitive to effort were heart period (HP) and blood pressure (BP). In the second and third experiment, the frequency and duration of eye blinks were found to be sensitive to visual workload. Subjects blinked less when the visual information load was high, and when they blinked, the duration of a blink was short. Eye blinks were not influenced by mental effort. In these studies, however, the algorithm used to compute the duration was different from that used in the previous studies. In the present experiment, the two algorithms will be compared.

In the previous experiments, HP decreases and BP increases with increasing task demands. Although heart rate variability (HRV) has been found to be a sensitive index of mental workload in several studies (e.g. Mulder, 1980; Aasman, Mulder & Mulder, 1986; Backs &

Seljos, 1994), it was only sensitive in the first study. The cardiovascular control system is rather complex. In order to know what effect effort has on this system, the working of the short term cardiovascular control system will be briefly outlined. The reason why HP, BP and HRV are affected by effort is that effort influences the autonomic nervous system. The heart is controlled by the parasympathetic and sympathetic branches of the autonomic nervous system. These branches mostly work reciprocally; when the parasympathetic activity increases, the sympathetic activity decreases and *vice versa*. An increase in parasympathetic activity and decrease in sympathetic activity leads to an increase in HP, while BP is only increased when the sympathetic activity increases. The parasympathetic and sympathetic systems also influences the variability of the heart beats (HRV). When the parasympathetic activity decreases and the sympathetic activity increases, the heart beats will become more regular. When there is no autonomic activity, the heart will beat like a metronome. HRV can be calculated by means of spectral analysis. In this way the variability within specific frequency bands can be obtained. The spectrum is mostly divided into three spectral bands: 0.02−0.06 (low-band), 0.07−0.14 (mid-band) and 0.15−0.50Hz (high-band). The parasympathetic system reacts much faster than the sympathetic system. Therefore, parasympathetic activity is reflected in all spectral bands, while sympathetic activity is only reflected in the two lower frequency bands. By comparing the HRV between the bands, information can be obtained about the parasympathetic and sympathetic influences, which is not possible for HP.

In previous studies, is has been found that HRV was not only influenced by effort, but also by respiration. HRV increased when the respiratory rate decreased and/or the respiratory volume increased. This happened often during high workload in which case the effect is opposite to that of effort. Thus respiration can disturb the effect on HRV. It has also been found that a better way to look at HRV is to divide the HRV by the variability in BP (BPV). This index is called the modulus and expresses the change in HR as the BP changes one mmHg. Because effort influences the autonomic system, changes in BP are reflected less by changes in HR. Respiration has almost the same effect on BPV as on HRV. Therefore, the modulus is virtually unaffected by respiration. The modulus can also be calculated by means of spectral analysis, separately for the three frequency bands (Mulder, 1988). For the modulus, like for HRV, parasympathetic activity is reflected in all bands while sympathetic activity is only reflected in the two lower bands.

There are far more factors influencing the heart than described here. Therefore, more information can be obtained about how effort influences the heart beat when several cardiovascular parameters are calculated than when only HP is used. When HP, BP, HRV and the modulus are used, more information about the way effort influences the autonomic nervous system can be obtained.

In the present experiment pilots had to fly in the simulator of the Human Factors Research Institute. There were two types of tasks: pursuit and tunnel task. In the pursuit task, subjects had to follow a target at a large distance. In the tunnel task, they had to fly trough a tunnel and perform simultaneously an auditory memory task with four levels of difficulty. Before and after the experiment, physiological recordings were obtained while the pilots sat quietly

in the simulator (rest). Physiological reactions were compared between rest and flight, between pursuit- and tunnel task and between the four levels of the tunnel task.

The motivation to invest effort was expected to be high because pilots were tested in pairs. When one was executing the task, the other observed. The pursuit task requires less effort than the tunnel task because no memory task has to be performed and the subject can stay at a comfortable distance behind the target jet. It is expected that the more difficult the tunnel task is, the more effort will be invested.

2 METHOD

2.1 Subjects

Twelve subjects (nine males and three females) participated in the experiment. The mean age was 30 (between 22 and 43). All subjects had flight training. Seven subjects were helicopter pilots (including the three females), two were F16 fighter jet pilots, two subjects had a private flight license and one subject was a helicopter navigator.

2.2 Simulator

The task was presented in a simulator, which has a spherical dome (radius 3 m; projected image of 156° horizontal and 42° vertical). The projected data base (desert environment) was generated with a graphics system with three parallel channels (Evans & Sutherland, ESIG-2000). A target jet (F-18) flew over the desert scene. The subjects sat in a mock-up of a cockpit with a force stick on the right and a throttle on the left side. On the stick was a button for the memory task (see § 2.3).

A display with a synthetic picture of the world (60° vertical and 45° horizontal), generated by an IRIS graphical computer, was positioned in front of the subject. System information (airspeed, altitude, heading and rate of descent) was presented on the display and in a head up display (HUD) that was projected on the dome. The HUD also contained a pitch ladder.
As can be seen in Fig. 1, the tunnel consisted of two series of squares (200×200 ft) connected by lines. The little squares (50×50 ft) indicated the centre of the tunnel. The target jet (indicated by the cross symbol) flew through the upper edge of the tunnel. The position of the target jet was also presented on a radar that could be used when the target jet was behind of the subject. The distance to the target jet was indicated at the left of the radar. The centre of this indicator corresponded to the required follow distance.

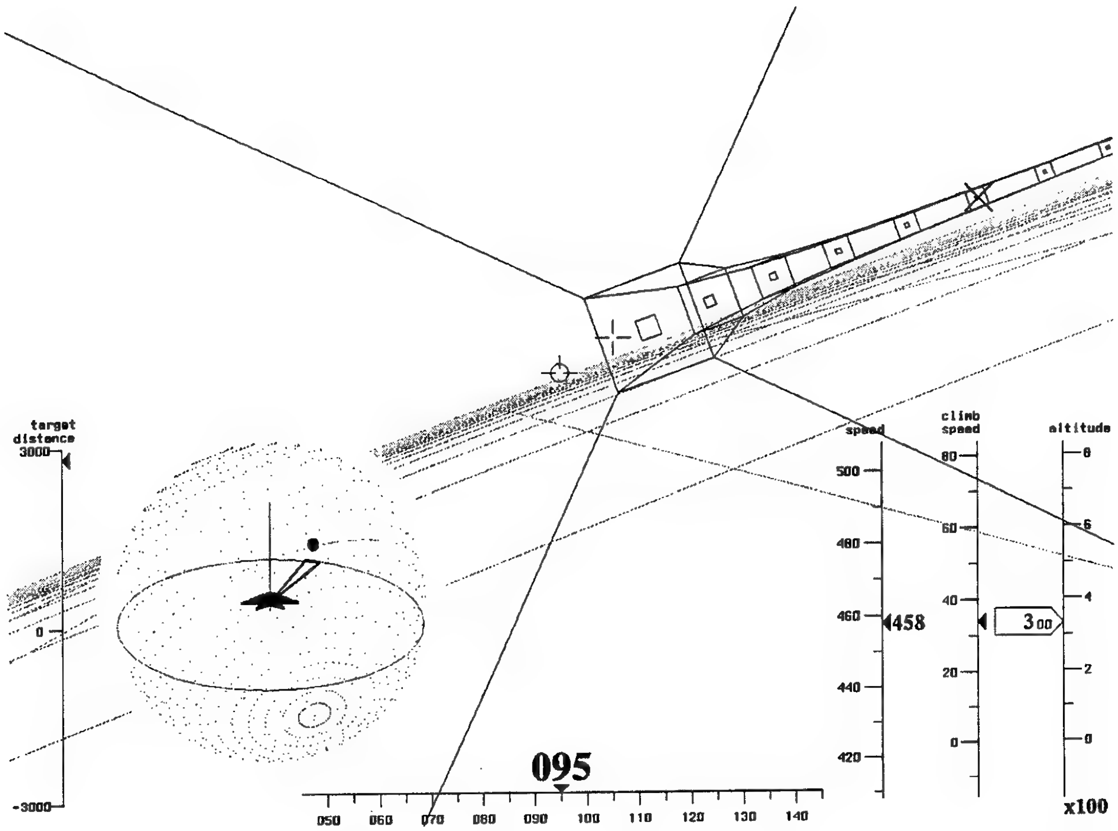


Fig. 1 Overview of the display with tunnel.

2.3 Tasks

The subjects had to do two tasks: a pursuit and a tunnel task. In the pursuit task, subjects had to follow a target jet at a large distance. This task was expected to be easy because the large distance enabled a thorough anticipation to the manoeuvres of the target jet. In the tunnel task, subjects had to fly through a tunnel and simultaneously perform a memory task. The tunnel task was presented on a display in front of the subject. The subjects were asked to fly as accurate as possible through the centre of the tunnel and stay at a fixed distance (1000 ft) behind a target jet that flew also through the tunnel. There were four levels of task difficulty (T1 to T4). The curves in the horizontal plane varied from smooth in T1 to sharp in T4. The curves in the vertical plane were varied from no variation in T1 to much in T4. Since the target jet flew always with a fixed speed of 500 kts, more power (thrust) was needed when the vertical slope of the tunnel was positive and less thrust was needed when the slope of the tunnel was negative. Keeping the target jet at a fixed distance was rather difficult due to variations in the vertical slope.

In the memory task (CMT) subjects had to respond to auditory target letters by pressing a button on the stick. The target letters were A, AB, ABC or ABCD (T1 to T4). Besides responding to the target letters, the subjects had to count the letters separately. The button had to be pressed twice when one of the letters appeared for the third time. Feedback was given when the subjects pressed twice after the third presentation of a letter ("GOOD"), when

the subject omitted to press twice or when he/she pressed twice at the wrong moment ("WRONG"). The tally of the last presented target letter had to be set to zero after feedback was given ("GOOD" or "WRONG"). The average time in between two letters was 3 ($\pm$0.5 s). 30% of the letters were targets that were never presented successively.

2.4 Performance and subjective ratings

Performance

In the tunnel task, flight performance was measured by computing the root mean square (RMS) of the deviation from the centre line of the tunnel and the RMS of the deviation from the required follow distance. The performance on the CMT was measured by the reaction times (RT) and the percentage of correct responses to targets that appeared for the third time. Performance in the pursuit task was not analysed.

Rating scale

Subjective effort was measured with the Rating Scale Mental Effort (RSME; Zijlstra, 1993). This scale was presented at the display at the end of each task. The subjects indicated how much effort they had invested during the previous block by moving an arrow along a vertical scale with a mouse. The RSME ranges from 0 to 150 and has labels ranging from "not effortful" (score 3) to "awfully effortful" (score 114).

2.5 Physiological measures

The electro-cardiogram (ECG), respiration, BP and electro-oculogram (EOG) were digitally recorded with the CODAS system (DATQ Instruments Inc.). The sample rate was 200 Hz for each channel. Pre-processing of the signals was done with the CODAS software. Spectral analysis and the calculation of mean values were done by Carspan (Mulder, 1988).

ECG

The times of the R-peaks from the ECG were automatically detected. Omissions were corrected manually after visual inspection of the data. The HP was calculated from these R-peak times.

Blood pressure

BP was recorded continuously with the TNO Portapres tonometer that measures both the systolic and the diastolic BP from beat to beat. The system has two cuffs; one was placed around the middle finger and one around the ring finger. After fifteen minutes the measurement switched to the other finger. Subjects were instructed not to bend the two fingers. Systolic and diastolic values from the blood pressure signal were detected automatically. Omissions were corrected manually. Values were excluded when the apparatus was being

calibrated or when the signal was disturbed. Most artifacts were caused by bendings of the fingers. Missing values were replaced by linearly interpolated values by the Carspan program. Interpolations were not conducted when there were too many missing values in succession (maximum 10). These parts were not included in further analysis.

Respiration

Respiration was measured by means of inductive plethysmography (Respitrace Inc.). The expansion of the chest and abdomen was measured with two elastic belts. The two signals were averaged. Peaks and valleys from this averaged signal were detected by CODAS and used for the calculation of the following respiratory parameters: total cycle times (T_{tot}), amplitude (ampl), inspiratory time (T_i), inspiratory flow (If) and duty cycle time (DCT). Fig. 2 gives a schematic view of a respiratory cycle with the parameters. Because respiration was not calibrated, no volume could be distracted from the signals. All conditions were flown in succession while the subjects were sitting in a stable position, which makes it unlikely that the amplitude changes due to movements of the belts. Therefore, the amplitude of the signal could be used instead of volume. Besides this analysis in the time domain, respiration was also analysed in the frequency domain (see analysis for more details).

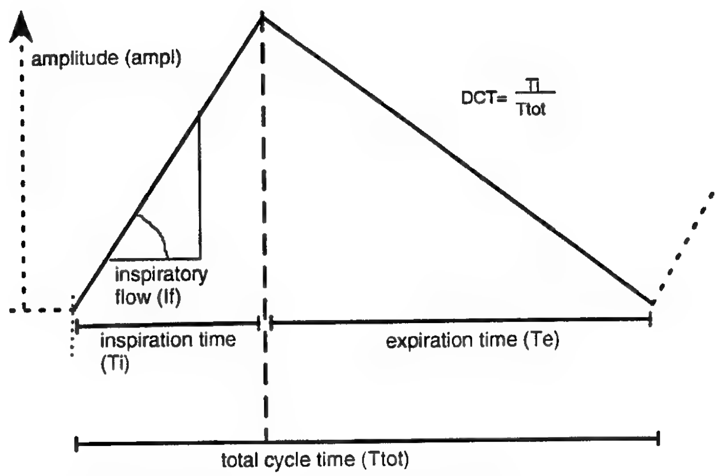


Fig. 2 Overview of a respiratory cycle with the parameters used in the present study.

Eye blinks

Eye blinks were derived from the EOG, measured with electrodes above and below the left eye. The EOG was recorded with an AC coupled amplifier (τ=10 s). The following parameters were distracted from the signal: number of blinks per minute (blink frequency), duration of a blink and blink amplitude. The duration of a blink was derived with two methods. The first method was the same as in the previous experiments (Veltman & Gaillard, 1994; Veltman, Gaillard & Van Breda, 1995), in which the derivative of the EOG was calculated. The duration of a blink was defined as the time between a valley and a peak in this new signal. In the second method the point halfway the amplitude of the first EOG flank was determined. The distance from this point to the point at the same level at the

second flank was defined as the blink duration (see Fig. 3 for more details). This method is used more often in the literature (e.g. Goldstein, Bauer & Stern, 1992; Wilson & Fullenkamp, 1989). In this study the two methods will be compared.

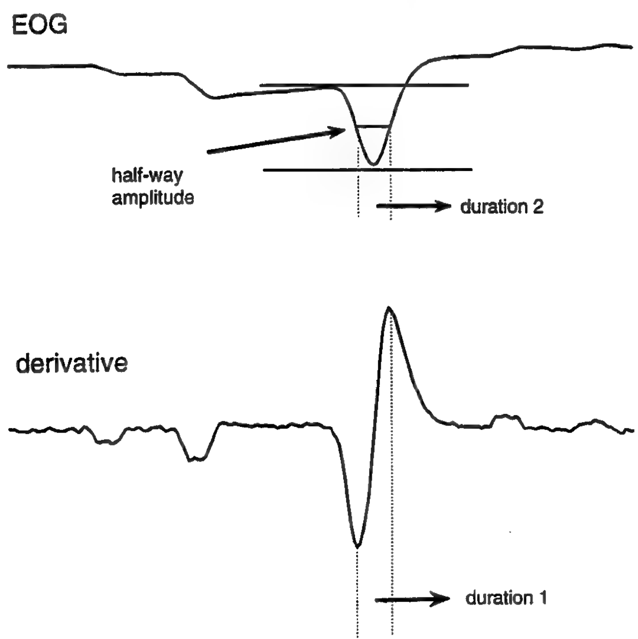


Fig. 3 EOG signal and its derivative. The duration of a blink is derived from the EOG signal at the half-way amplitude and from the derivative from valley to peak.

2.6 Procedure

Subjects participated in the experiment for two days. Two subjects were alternately tested. The first afternoon they were trained to fly through the tunnel and to do the CMT. They first flew four times through the tunnel with different levels of difficulty without CMT and then with the CMT, starting with one target end ending with four targets. The experiment took place on the second afternoon. Fig. 4 gives a schematic overview of the conditions.

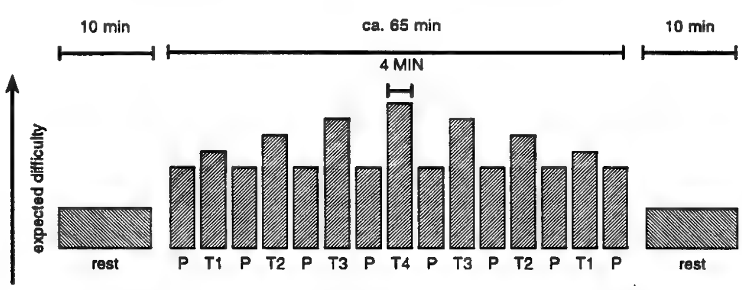


Fig. 4 Schematic representation of the task blocks (see text for explanation).

Subjects started with a rest of 10 minutes in which they were asked to sit in the simulator with the eyes open. After the rest they followed 15 times a target jet, for four minutes. During the baseline blocks (P) they followed the target jet without any restrictions. During the other blocks (T1 to T4) they had to fly through a tunnel and had to perform a CMT. The difficulty of the tunnel varied from relatively easy (T1) to difficult (T4). The difficulty of the CMT was varied in the same way as the difficulty of the tunnel.

2.7 Analysis

Spectral analysis

Spectral analysis was done for HR, systolic BP, diastolic BP and respiration with the software package Carspan (Mulder, 1988). The spectral energy in the mid-band (0.07–0.14 Hz) and the high-band (0.15–0.50 Hz) was calculated within 40 s windows that shifted with 10s steps. Thus spectral values were available for each 10 s. Thus there was 75% overlap between the adjacent windows. The spectral values were divided by the squared means in the 40 s window to make the variability statistically independent of the mean HR.
Coherence and modulus were calculated between systolic BP and HP. The coherence expresses the resemblance of two signals in the frequency domain and is similar to a squared correlation in the time domain. The modulus expresses the gain between two signals. Since there will always be a gain between two signals, even if there is no functional relation, the average modulus within a frequency band is based on frequencies for which the coherence is greater than 0.5.

Statistical analysis

The results were analysed with repeated measures analysis of variance conducted by the software package Statistica. The multivariate approach (MANOVA) was used when a factor had more than two levels, because this approach is more robust to violations of the assumptions (compound symmetry and sphericity) for repeated measures analysis of variance than the univariate approach (ANOVA) (Statistica, 1994). For the MANOVA's, the Wilks' λ are reported and for the ANOVA's the F ratios.
The performance data includes 7 tunnel blocks and the physiological data of 17 blocks (2 rest+7 tunnels+8 pursuit) which is too much for a complete analysis. Therefore, the analysis for the physiological data is split up in two parts:

- differences between rest, pursuit and tunnel, in which the data for the 8 pursuit segments and 7 tunnels are averaged;
- differences among the seven levels of the tunnel. Significant main effect were analysed further by the following planned comparisons (see also Table I):
  1 the first part of the experiment and the second part
  2 T1 *versus* T2
  3 T1 *versus* T3
  4 T1 *versus* T4
  5 T2 *versus* T3
  6 T2 *versus* T4
  7 T3 *versus* T4.

Table I Contrasts for the planned comparisons. Tunnels with the same sign are taken together and positive signs are tested against negative signs. The sum of the contrasts must always be zero which explains why tunnel 4 has a value of −2.

		tunnel						
		first part				second part		
	analysis	1	2	3	4	3	2	1
1	first *vs* second part	1	1	1		−1	−1	−1
2	T1 *vs* T2	1	−1				−1	1
3	T1 *vs* T3	1		−1		−1		1
4	T1 *vs* T4	1			−2			1
5	T2 *vs* T3		1	−1		−1	1	
6	T2 *vs* T4		1		−2		1	
7	T3 *vs* T4			1	−2	1		

3 RESULTS

3.1 Performance

RMS

No significant differences in the RMS of the deviation from the centre of the tunnel and the RMS of the deviation of the follow distance were found among the tunnel blocks. There were large individual differences in RMS scores. This was also reflected by the percentage of time that was flown outside the tunnel which ranged from 2% to 38%. These percentages did also not differ among the tunnel blocks.

CMT

The RT differed among the 7 tunnel blocks [Wilks' λ=0.017, p<0.001]. RT increased when the size of the target set increased (see Table II). The percentage correct responses did not differ significantly (see Fig. 5). The planned comparisons on the RTs showed that there was no difference between the first and second half of the experiment.

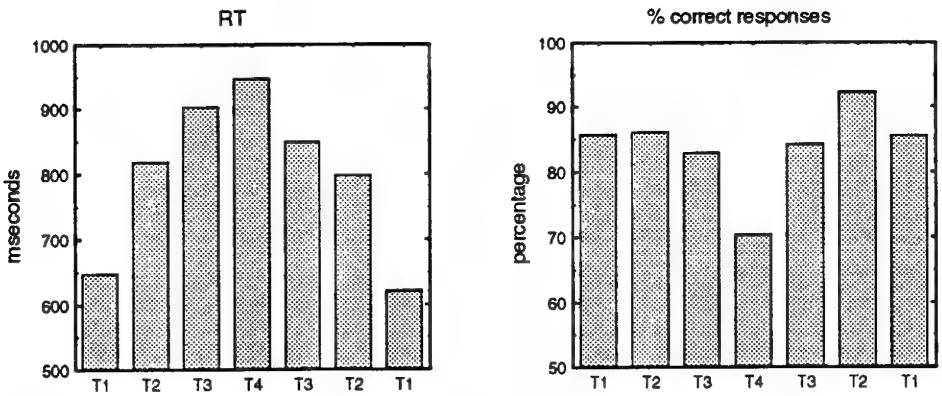


Fig. 5 Reaction times and percentage of correct responses to targets that appeared for the third time.

Table II Results of the contrast analysis of the RT.

	$F(1,11)$
first *vs* second half	2
T1 *vs* T2	41***
T1 *vs* T3	195***
T1 *vs* T4	214***
T2 *vs* T3	15 **
T2 *vs* T4	70***
T3 *vs* T4	22***

*** $p<0.001$
** $p<0.01$

3.2 Subjective ratings

The tunnel task is rated as more effortful than the pursuit task [score 61 and 24, respectively; $F(1,11)=69.7$, $p<0.001$]. The ratings on the different levels in the tunnel task are presented in Fig. 6. The levels of the tunnel were rated significantly different (Wilks' $\lambda=0.07$, $p<0.01$).

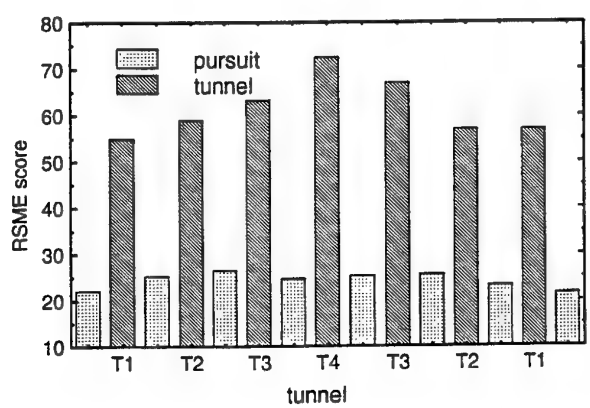


Fig. 6 RSME scores for the levels of the tunnel.

The planned comparisons revealed that: T3 and T4 were more difficult than T1 [$F(1,11)=12.5$, $p<0.01$; $F(1,11)=23.1$, $p<0.001$, respectively], T3 and T4 were more difficult than T2 [$F(1,11)=7.8$, $p<0.05$; $F(1,11)=23.1$, $p<0.001$, respectively] and T4 was more difficult than T3 [$F(1,11)=7.3$, $p<0.05$].

3.3 Physiological measures

The physiological measures are presented in Figs 7, 8 and 9. They are tested in two ways, 1) differences between rest, pursuit and tunnel task and 2) differences between levels of difficulty in the tunnel task. The results of the analyses are presented in Tables III and IV. The first analysis showed significant main effects for all measures, except DCT. The planned comparisons showed significant effects among rest, pursuit and tunnel task for all cardiovascular measures. The planned comparisons for respiratory measures (except DCT) showed differences between rest and pursuit and between rest and tunnel. The difference between pursuit and tunnel is only significant for CT and Ti.

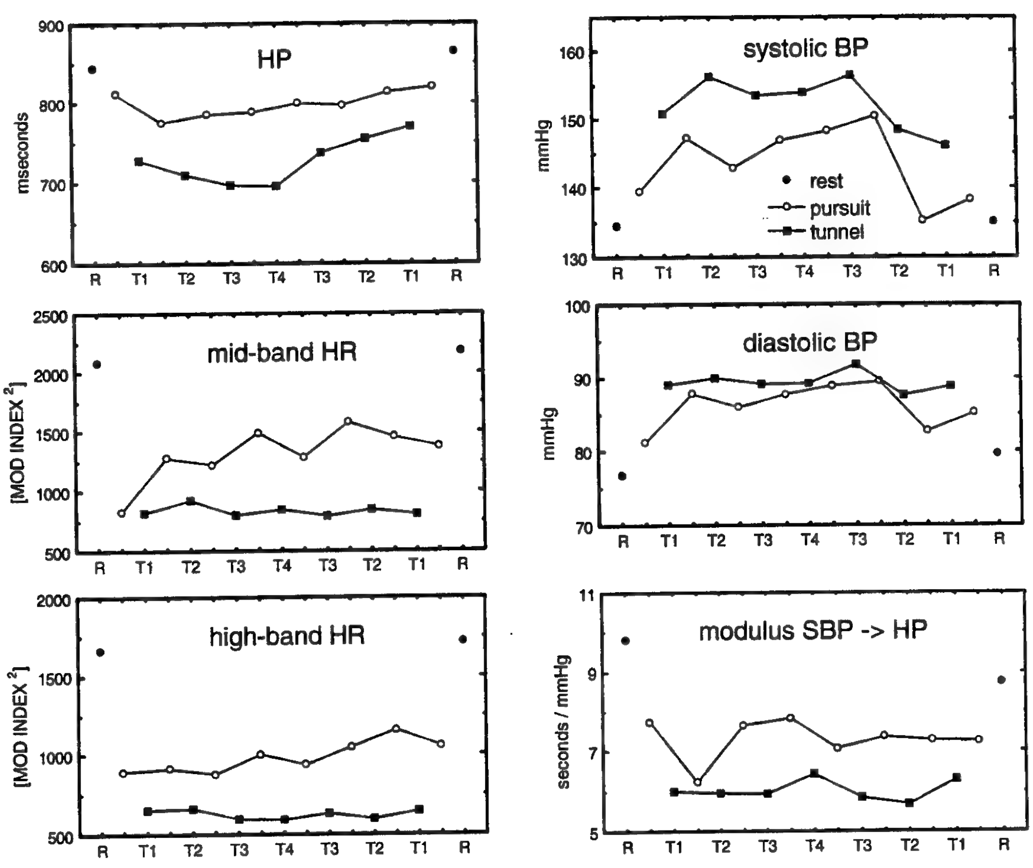


Fig. 7 Cardiovascular results.

The eye blink parameters all show significant main effects. Planned comparisons revealed that blink interval was significant for all comparisons and that amplitude differed only for

pursuit *versus* tunnel. The two metrics for the duration show rather different results. The first method that was also used in the previous experiment, showed only differences between rest and the two task conditions and the second method showed differences between rest *versus* tunnel and pursuit *versus* tunnel.

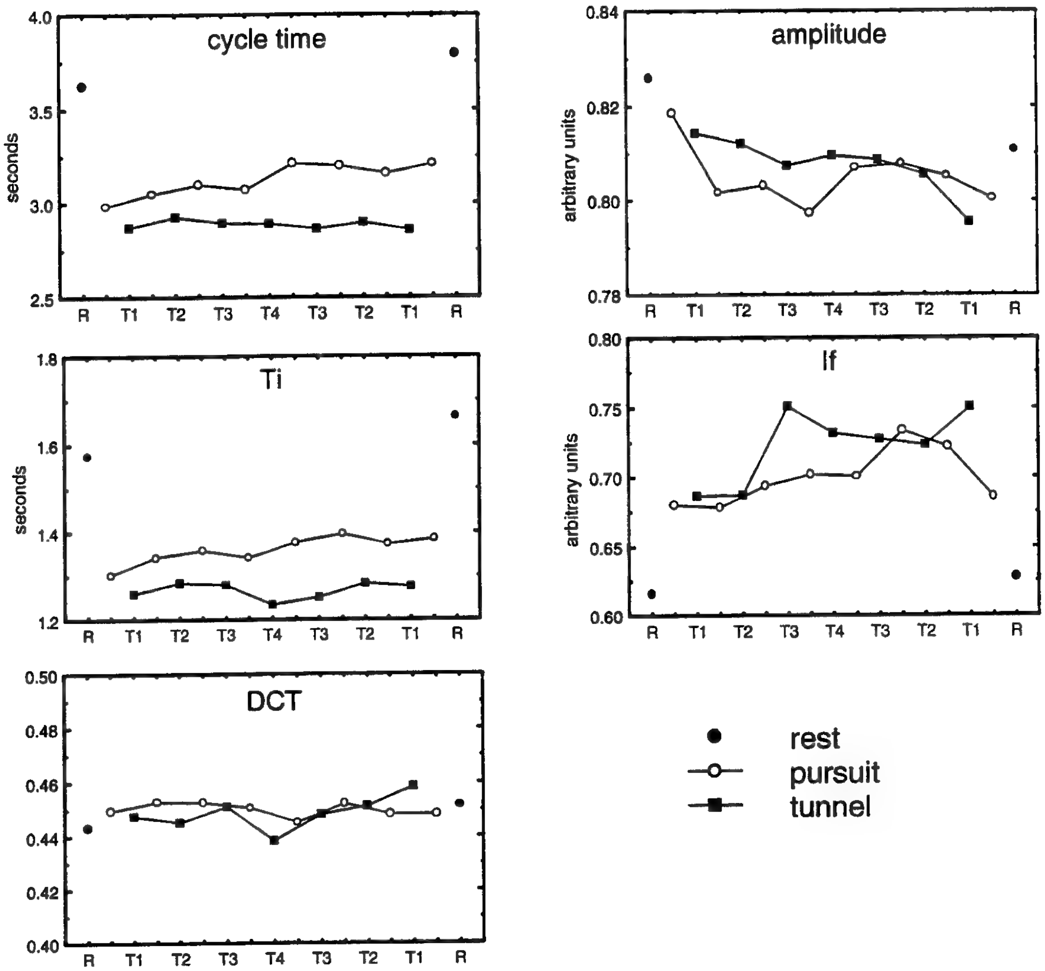


Fig. 8 Respiratory results.

Table IV shows that HP differed significantly among all levels of the tunnel and systolic BP showed a significant difference between T1 and T3. The other cardiovascular measures showed no differences. Amplitude is the only respiratory measure that showed a significant main effect. However, planned comparisons between the levels of the tunnel showed no significant effects. This is possible because not all possible combination are tested in the planned comparisons. In addition, the main effect is tested by means of a MANOVA and the planned comparisons by means of an ANOVA which might reveal differences.

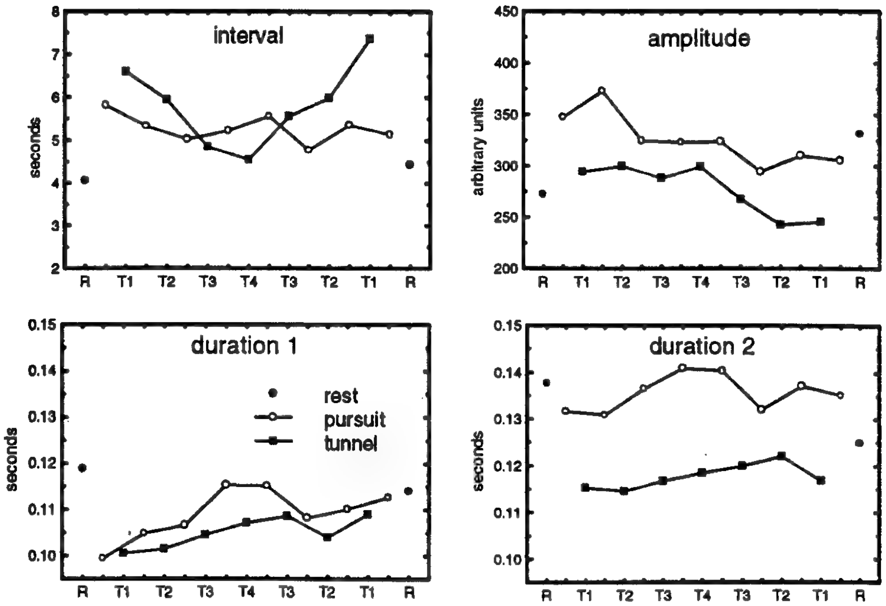


Fig. 9 Eye blink results.

Blink interval and amplitude showed significant main effects. However, only blink interval showed significant effects in the planned comparisons. It decreases when difficulty increases.

Table III Statistical results of the cardiovascular, respiratory and eye blink data: differences between rest, pursuit and tunnel.

	main effect	planned comparisons $F(1,11)$		
	Wilks' λ	rest *vs* pursuit	rest *vs* tunnel	pursuit *vs* tunnel
cardiovascular				
HP	0.18***	20.4***	50.8***	32.0***
sys. BP	0.28***	17.3**	27.7***	19.6**
dias BP	0.18***	37.2***	50.3***	7.6*
mid-band HR	0.42*	5.6*	13.3**	6.8*
high-band HR	0.25***	15.4**	29.2***	17.1**
modulus	0.26**	16.1**	28.6***	7.8**
respiration				
CT	0.20***	27.6***	44.4***	13.8**
Ti	0.21***	21.3***	23.7***	17.8**
DCT	0.99	– –	– –	– –
amplitude	0.49*	11.2**	5.5*	ns
If	0.48*	11.9**	8.4*	ns
eye blinks				
interval	0.50*	6.75*	11.2**	6.3*
duration 1	0.25**	8.8*	30.6***	ns
duration 2	0.17***	ns	13.3**	11.8**
amplitude	0.55*	ns	ns	4.9*

*** $p<0.001$ ns   not significant
** $p<0.01$ – –   not analysed because of insignificant main effect
* $p<0.05$

Table IV Statistical results of the cardiovascular, respiratory and eye blink data: differences between the levels of the tunnel.

	main effect	planned comparisons $F(1,11)$						
	Wilks' λ	first part *vs* second part	T1 *vs* T2	T1 *vs* T3	T1 *vs* T4	T2 *vs* T3	T2 *vs* T4	T3 *vs* T4
cardiovascular								
HP	0.09**	23.6***	13.9**	17.2**	31.6**	8.8*	37.6***	14.7**
sys. BP	0.11*	ns	ns	11.2**	ns	ns	ns	ns
dias BP	ns	--	--	--	--	--	--	--
mid-band HR	ns	--	--	--	--	--	--	--
high-band HR	ns	--	--	--	--	--	--	--
modulus	ns	--	--	--	--	--	--	--
respiration								
CT	ns	--	--	--	--	--	--	--
Ti	ns	--	--	--	--	--	--	--
DCT	ns	--	--	--	--	--	--	--
amplitude	0.14*	ns	ns	ns	ns	ns	ns	ns
If	ns	--	--	--	--	--	--	--
eye blinks								
interval	0.14*	ns	5.9*	7.7*	46.0***	ns	16.8**	7.9*
duration 1	ns	--	--	--	--	--	--	--
duration 2	ns	--	--	--	--	--	--	--
amplitude	0.13*	ns	ns	ns	ns	ns	ns	ns

*** $p<0.001$ ns not significant
** $p<0.01$ -- not analysed because of insignificant main effect
* $p<0.05$

Correlations were calculated between HRV in the mid-band, BPV in the mid-band, the modulus, respiratory cycle time and amplitude to explore the influence of respiration upon HRV. Only rest periods are included in the analysis because task periods are confounded with the task effects. For example HRV and CT can decrease when more effort is invested due to increasing task difficulty. A positive correlation is then partly caused by the tasks and not only because respiration influences HRV. The data for the correlations consisted of values within the 40 s window. For each window, one value of HRV, BPV and the modulus is available. CT and VT are average values within each window. HRV and BPV in the mid-band, CT and amplitude are standardized (z-transformations) within each subject to correct for individual differences. Table V shows that the highest correlations are found between the respiratory parameters and HRV. CT does not correlate with BPV. Respiratory amplitude correlates with HRV and BPV and not with the modulus.

Table V Correlations between HRV in the mid-band, modulus, and respiratory cycle time and amplitude (n=1277).

	HRV mid-band	BPV mid-band	modulus
CT	.21***	$.02^{ns}$	.15***
amplitude	.34***	.17***	$.02^{ns}$

*** : $p<0.001$ ns: not significant

4 DISCUSSION

Large changes

Mental effort was manipulated at two levels in the present study: *large* changes in effort between rest, pursuit and tunnel task, and *small* changes within the different levels of difficulty in the tunnel task. All physiological measures obtained in the present experiment were sensitive to large changes in effort. Differences between the tunnels were only found for HP, systolic BP and blink interval, although systolic BP differed only between T1 and T3 and not between T1 and T4 as would be not expected. HP showed systematic differences between all levels of the tunnel; the difference between T1 and T2 was smaller than between T1 and T3 etc. Thus, it can be concluded that HP is sensitive to small changes in effort.

Small changes

HRV and the modulus between systolic BP and HP were not sensitive to the differences between the levels of the tunnel. The effects of rest, pursuit and tunnels were more stable for the modulus than for HRV. It is likely that this is caused by respiration. The correlations between respiration and HRV in the mid-band and the modulus during rests showed that respiration influences HRV more than the modulus. Subjects breathe deeply and long from time to time during task performance. This will increase the HRV more than the modulus and thus HRV will fluctuate more than the modulus. This corresponds with the results of the previous experiment (Veltman *et al.*, 1995) in which HRV was found to be influenced more by respiration than by the modulus. The modulus is a more direct indication of the suppression of the cardiovascular control system than the HRV (Mulder, 1988; Steptoe, Fieldman & Evans, 1993). Thus, the modulus is a better index of mental effort than HRV.

Respiration

From respiratory parameters only CT and Ti showed stable effects among rest, pursuit and tunnel task. However, no differences were found among different levels of the tunnel task. Because Ti is part of CT and does not show stronger effects than CT, it does not provide additional information. Thus, CT appears to be the best respiratory measure for mental effort.

Eye blinks

Blink interval increased during flight compared to rest. During the easy levels, blink interval is longer than during the difficult levels of the tunnel task. Because the visual information to be processed is assumed to be the same for each level, it can be concluded that blink interval decreases as mental load increases. In the previous studies (Veltman *et al.*, 1994, 1995), blink interval was found to be independent of mental load. In that study, no differences were found between flying with or without an CMT. However, in those studies, there were no differences in visual load between the condition with and without the CMT, while the visual load in the tunnel task is much higher than in the pursuit task. Furthermore, the memory set size in the previous studies was always 4. Thus, it seems that blink interval increases when visual load increases, but that it is decreased with a higher memory load. The last effect has also been found by De Jong and Merkelbach (1990), who found a relation between rehearsal and decrease in blink interval. They showed slides to subjects and asked them to rehearse that information. When the slides were presented, blink interval increased and with rehearsal it decreased. Other researchers have found an increase in blink interval due to cognitive activity like arithmetic (*e.g.* Stern, Walrath & Goldstein, 1984). De Jong and Merkelbach hypothesised that the difference between arithmetic and rehearsal is that during rehearsal subjects speak subvocally. The representation of the eyelid and eyeball in the primary motor cortex is bordered by that of the tongue, larynx and face. Stimulating the muscles in the mouth, as is the case with rehearsal, probably also stimulates the muscles in the eyelid, resulting in a decreased blink interval during rehearsal. This can also explain the results found in the present experiment. It is reasonable that subjects rehearsed more when they had to count four letters than when they had to count only one letter.

Determination of duration

The blink duration derived with the method in the previous experiments (duration 1) was rather different from duration 2 that was used by other researchers (*e.g.* Bauer, Goldstein & Stern, 1987; Wilson & Fullenkamp, 1989; Stern, Boyer & Schroeder, 1994). Duration 1 showed no significant difference between pursuit and tunnel, while duration 2 showed a large difference. In the two previous experiments, duration 1 also showed no differences between flight with or without a CMT. The time between the maximum slope in the downstroke and upstroke does not change due to a memory load, while the time between the halfway amplitude increases due to memory load. Because duration 1 is an indirect measure (it is derived from the derivative of the EOG signal) it is preferable to use duration 2. Furthermore, duration 2 is also used by other researchers which makes data comparisons possible. The effect of rehearsal is not found for blink duration. An increase in visual load leads to an increase in blink interval and a decrease in blink duration. Rehearsal will only affect blink interval. The effects of visual load and rehearsal can be separated when both blink interval and duration are measured.

Effort

The aim of the present experiment is to examine whether the physiological measures would be sensitive to different levels of difficulty in the tunnel task. For this, it was important that

the subjects were motivated and that the effort differed among the levels of difficulty in the tunnel task. There are no direct ways of measuring motivation and effort and thus it has to be deduced from the performance and subjective ratings. No differences among the levels of the tunnel were found in flight performance. It can be concluded from this that subjects invested more effort to keep the performance at an acceptable level as the task became more difficult. Furthermore, no differences were found in the percentage of correct responses in the CMT task. More errors were expected when subjects did not increase their effort as the memory set increased. The RTs increased as the set size increased. This is not because subjects invested less effort but because more comparisons have to be made in memory in order to decide whether a letter was a target letter. The subjective ratings indicated that subjects invested more effort as the level of difficulty increased. This is no guarantee that subjects indeed increased their effort since subjects have difficulties in making a distinction between all the aspects of workload (O'Donnell & Eggemeier, 1986). It was obvious for the subjects that the difficulty increased. Thus, it is possible that the subjects rated the task demands and not their effort. However, the difference between task demand and effort was made clear to the subjects. Furthermore, the performance of a subject could be observed by his/her colleague, which forced the subject to do his best. Therefore, we are rather sure that the ratings reflected the effort. When the subjective ratings and the performance are taken together, it can be concluded that the effort increased as the task became more difficult. Thus, most physiological measures are not sensitive to small changes in effort.

Adaptation

The physiological system changes because subjects have to adapt their physiological system to the changing task demands (Hockey, 1986). It is possible that differences within tasks require less adaptation than differences between tasks. In other words, the preparation to perform an easy tunnel task with one target letter in the CMT is almost the same as the preparation for a difficult tunnel task with 4 target letters, because the type of tasks are the same. This should be explored in future studies. If this hypothesis holds, physiological measures will be more sensitive when both the type of task and the difficulty changes.

5 CONCLUSIONS

Physiological measures

- Heart period, heart rate variability, blood pressure and respiration reflected large differences in mental effort.
- Heart period is sensitive to small changes of mental effort.
- Respiratory cycle time was more sensitive to mental effort than the other respiratory parameters.
- The modulus between systolic blood pressure and heart period is a more stable measure of mental effort than heart rate variability, since it is influenced less by respiration.

Eye blinks

- Blink interval increases and blink duration decreases due to visual load.
- Blink interval decreases with higher memory loads. Blink duration is not affected by memory load.
- Blink interval and duration together provide valuable information about the visual load.

REFERENCES

Aasman, J., Mulder, G. & Mulder, L.J.M. (1987). Operator effort and the measurement of heart rate variability. *Human Factors, 29(2),* 161-170.

Adams, M.J., Tenney, Y.J. & Pew, R.W. (1991). *Strategic workload and the cognitive management of advanced multi-task systems* (Report SOAR CSERIAC 91-6). Wright Patterson AFB, OH: Crew System Ergonomics, Information Analysis Center.

Backs, R.W. & Seljos, K.A. (1994). Metabolic and cardiorepiratory measures of mental effort: the effects of level of difficulty in a working memory task. *International Journal of Psychophysiology, 16,* 57-68.

Bauer, L.O., Goldstein, R. & Stern, J.A. (1987). Effects of information-processing demands on physiological response patterns. *Human Factors, 29(2),* 213-234.

De Jong, P.J. & Merkelbach, H. (1990). Eyeblink frequency, rehearsal activity, and sympathethic arousal. *International Journal of Neuroscience, 51,* 89-94.

Gaillard, A.W.K. & Wientjes, C.J.E. (1994). Mental load and work stress as two types of energy mobilization. *Work & Stress, 8,* 141-152.

Goldstein, R., Bauer, L.O. & Stern, J.A. (1992). Effect of task difficulty and interstimulus interval on blink parameters. *International Journal of Psychophysiology, 13,* 111-117.

Hart, S.G. (1989). Crew workload-management strategies: A critical factor in system performance. *Proceedings of the fifth International Symposion on Aviation Psychology* Columbus, OH: 22-27.

Hart, S.G. & Staveland L.E. (1988). Devellopement of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P.A. Hancock & N. Meshkati (Eds.), *Human Mental Workload* (pp. 139-184). Amsterdam: Elsevier Science Publishers BV.

Hendy, K.C., Hamilton, K.M. & Landry, L.N. (1993). Measuring subjective workload: When is one scale better than many? *Human Factors, 35(4),* 579-601.

Hockey, G.R.J. (1986). A state control theory of adaptation to stress and individual differences in stress management. In G.R.J. Hockey, A.W.K. Gaillard & M.G.H. Coles (Eds.), *Energetics and human information processing* (pp. 285-298). Dordrecht: M. Nijhoff Publishers.

Hockey, G.R.J. (1993). Cognitive energetic control mechanisms in the management of work demands and psychological health. In A.D. Bradley & L. Weiskrantz (Eds.), *Attention, Selection, Awareness and control: A tribute to Donald Broadbent* (pp. 328-345). Oxford University Press.

Kantowitz, B.H. (1988). Mental Workload. In P.A. Hancock (Ed.), *Human Factors Psychology* (pp. 81-119). Amsterdam: Elsevier Science Publishers BV.

Mulder, G. (1980). *The heart of mental effort.* Thesis. Groningen: University of Groningen.

Mulder, L.J.M. (1988). *Assessment of cardiovascular reactivity by means of spectral analysis.* Thesis. Groningen: University of Groningen.

Norman, D.A. & Bobrov, D.G. (1975). On data-limited and resource-limited processes. *Cognitive Psychology, 7,* 44-64.

O'Donnell, R.D. & Eggemeier, F.T. (1986). Workload assessment methodology. In K. Boff, L. Kaufman & J.P. Thomas (Eds.), *Handbook of Perception and Human Performance* (pp. 42.1-42.49). New York: Wiley.

Statistica (1994). *Handbook.* Statsoft Inc.

Steptoe, A., Fieldman, G. & Evans, O. (1993). An experimental study of the effects of control over work pace on cardiovascular responsivity. *Journal of Psychophysiology, 7,* 290-300.

Stern, J.A., Walrath, L.C. & Goldstein, R. (1984). The endogenous eyeblink. *Psychophysiology, 21,* 22-33.

Stern, J.A., Boyer, D. & Schroeder, D. (1994). Blink rate: A possible measure of "fatigue". *Human Factors, 36(2),* 285-297.

Stone, G., Gulick, R.K. & Gabriel, R.F. (1987). *Use of task timeline analysis to assess crew workload. The practical assessment of pilot workload* (AGARDograph 282). Neuilly sur Seine: AGARD.

Veltman, J.A. (1991). *Mentale werkbelasting bij vliegers: theoretisch kader en literatuuroverzicht* [Mental workload among pilots: Theoretecial background and literature review] (Report IZF 1991 A-33). Soesterberg, The Netherlands: TNO Institute for Perception[1].

Veltman, J.A. & Gaillard, A.W.K. (1993). *Evaluation of subjective and physiological measurement techniques for pilot workload* (Report IZF 1993 A-5). Soesterberg, The Netherlands: TNO Institute for Perception.

Veltman, J.A. & Gaillard, A.W.K. (1994). *Evaluation of pilot workload in a simulated environment* (Report TNO-TM 1994 A-48). Soesterberg, The Netherlands: TNO Human Factors Research Institute.

Veltman, J.A., Gaillard, A.W.K. & Van Breda, L. (1995). *Physiological reactions to workload peaks.* (Report TNO-TM 1995 A-33). Soesterberg, The Netherlands: TNO Human Factors Research Institute.

Wilson, G.F. & Fullenkamp, P. (1989). A comparison of pilot and WSO workload during training missions using phychophysiological data. *Congress Proceedings of the Western European Association for Aviation Phychology,* September 1989.

Zijlstra, F.R.H. (1993). *Efficiency in Work Behaviour: A design approach for modern tools.* Thesis. Technical University of Delft.

Soesterberg, 21 June 1996

Drs. J.A. Veltman
(1st author, project manager)

[1] On January 1, 1994 the name "TNO Institute for Perception" has been changed to "TNO Human Factors Research Institute".

REPORT DOCUMENTATION PAGE

1. DEFENCE REPORT NUMBER (MOD-NL)	2. RECIPIENT'S ACCESSION NUMBER	3. PERFORMING ORGANIZATION REPORT NUMBER
RP 96-0162		TM-96-A026
4. PROJECT/TASK/WORK UNIT NO.	**5. CONTRACT NUMBER**	**6. REPORT DATE**
789.4	A91/KLu/314	21 June 1996
7. NUMBER OF PAGES	**8. NUMBER OF REFERENCES**	**9. TYPE OF REPORT AND DATES COVERED**
28	28	Final

10. TITLE AND SUBTITLE

Physiological workload reactions to increasing levels of task difficulty

11. AUTHOR(S)

J.A. Veltman and A.W.K. Gaillard

12. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

TNO Human Factors Research Institute
Kampweg 5
3769 DE SOESTERBERG

13. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Director of Airforce Research and Development
Binckhorstlaan 135
2516 BA DEN HAAG

14. SUPPLEMENTARY NOTES

15. ABSTRACT (MAXIMUM 200 WORDS, 1044 BYTE)

The sensitivity of physiological measures to mental workload has been investigated in a flight simulator. Twelve pilots had to fly through a tunnel with different levels of difficulty. Additionally, they had to perform a memory task with four levels of difficulty. The easiest memory task was combined with the easiest tunnel task and the most difficult memory task with the most difficult tunnel task. Between each tunnel session, subjects had to fly a pursuit task in which a target jet had to be followed at a large distance. Rest periods before and after the experiment were used as a baseline for the physiological measures. Mental workload was measured with heart period, continuous blood pressure, respiration and eye blinks. Heart rate variability, blood pressure variability and the gain between systolic blood pressure and heart period (modulus) was derived from the measures. All measures showed differences between rest and flight, and between the pursuit task and the tunnel task. Only heart period was sensitive to the different levels of the tunnel. Heart rate variability is affected by respiration, and therefore, can only be interpreted together with the respiratory data. The modulus was hardly influenced by respiration and therefore, appears to be a better measure than heart rate variability. From the respiratory parameters, the duration of a respiratory cycle was the most sensitive to changes in workload. The time in between two successive eye blinks (blink interval) increased and the blink duration decreased as more visual information had to be processed. Increasing the difficulty of the memory task leaded to a decrement in blink interval, probably caused by subvocal activity during the rehearsal of the letters. Blink duration was not influenced by the memory load.

16. DESCRIPTORS	IDENTIFIERS
Blood Pressure Eye Blinks Heart Rate Pilot Psychophysiology Respiration Workload	

17a. SECURITY CLASSIFICATION (OF REPORT)	17b. SECURITY CLASSIFICATION (OF PAGE)	17c. SECURITY CLASSIFICATION (OF ABSTRACT)
18. DISTRIBUTION/AVAILABILITY STATEMENT		**17d. SECURITY CLASSIFICATION (OF TITLES)**
Unlimited availability		

VERZENDLIJST

1. Directeur M&P DO

2. Directie Wetenschappelijk Onderzoek en Ontwikkeling Defensie

3. { Hoofd Wetenschappelijk Onderzoek KL
   Plv. Hoofd Wetenschappelijk Onderzoek KL

4. Hoofd Wetenschappelijk Onderzoek KLu

5. { Hoofd Wetenschappelijk Onderzoek KM
   Plv. Hoofd Wetenschappelijk Onderzoek KM

6, 7 en 8. Bibliotheek KMA, Breda

9 en 10. Drs. G.J.P. van den Elzen, Directie Personeel KLu, Afd. Gedragswetenschappen, Den Haag

Extra exemplaren van dit rapport kunnen worden aangevraagd door tussenkomst van de HWOs of de DWOO.